DOCUMENT RESUME

ED 054 591 EC 033 305

AUTHOR Ayllon, T.; Barnes, Jarvis

TITLE A Comparison Between Standard Instruction and Reinforcement Program for the Trainable Retarded. Research and Development Report, Volume IV, Number 3.

INSTITUTION Atlanta Public Schools, Ga. Communication Skills Labs.

SPONS AGENCY Bureau of Elementary and Secondary Education (DHEW/OE), Washington, D.C.

PUB DATE 70

NOTE 38p.

EDRS PRICE MF-$0.65 HC-$3.29

DESCRIPTORS *Academic Achievement; Behavior Change; *Exceptional Child Research; *Mentally Handicapped; *Motivation Techniques; Operant Conditioning; Positive Reinforcement; *Trainable Mentally Handicapped

ABSTRACT

Three studies were conducted to compare standard instruction with reinforcement programs for the trainable mentally retarded (TMR). In the first study, 24 TMRs were matched into two groups, with one group participating in a 6-week academic program having token reinforcement and the other group continuing in a standard teaching program. Statistical analysis of pretest and posttest scores on a standardized test revealed significant improvement (t equals 2.55, significant at .01 level) for the experimental group, while the control group showed no improvement during the 6 weeks. The second study was designed to determine the role of reinforcement on performance on the criterion test. Findings revealed that both TMR groups in the first study showed score improvement when tested with consequences of maximal reinforcement for correct performance. The third study involved a duplication of the second study, but with normal students. The results showed statistically significant increases in scores achieved with maximal reinforcement consequences for correct performance. (CB)

ED0 54591

N-LJ
EC

RESEARCH AND DEVELOPMENT REPORT

VOL. IV, NO 3 SUMMER, 1970

A COMPARISON BETWEEN STANDARD INSTRUCTION AND REINFORCEMENT PROGRAM FOR THE TRAINABLE RETARDED SUMMER, 1970

Atlanta Public Schools

Atlanta, Georgia

EC 033 305 E

R E S E A R C H A N D D E V E L O P M E N T R E P O R T

Vol. IV, No. 3 Summer, 1970

A COMPARISON BETWEEN STANDARD INSTRUCTION AND REINFORCEMENT PROGRAM FOR THE TRAINABLE RETARDED

Funded Under ESEA Title I, P.L. 89-10

Prepared by

Dr. T. Ayllon
Professor of Psychology and Special Education
Georgia State University

and

Dr. Jarvis Barnes
Assistant Superintendent
for Research and Development

Dr. John W. Letson
Superintendent

Atlanta Public Schools
224 Central Avenue, S. W.
Atlanta, Georgia

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

PREFACE

In 1970, three studies were funded under the Elementary and Secondary Education Act of 1965 (ESEA), Title I (summer appropriation), and subcontracted to Dr. T. Ayllon, Professor of Psychology and Special Education, Georgia State University. The first report, "Token Reinforcement and Academic Objectives with the Trainable Retarded," was prepared by Dr. Ayllon with the assistance of F. Gerald McCullen, Kathleen Kelley, and Thomas Schneider, Jr. The second report, "A Comparison Between Standard Instruction and Reinforcement Program for the Trainable Retarded," was prepared by Dr. Ayllon with the assistance of Kathleen Kelley, and F. Gerald McCullen. A third article is a "Design for a Nine-Month School-Wide Program of Token Reinforcement for the Trainable Mentally Retarded." Since the data from the first study were a determining factor in planning the second study and the results from studies one and two were used in writing study number three, it is suggested that all three publications be read in the proper sequence.

Jarvis Barnes
Assistant Superintendent
for Research and Development

TABLE OF CONTENTS

	Page
ABSTRACT	1
INTRODUCTION	2
EXPERIMENT I	
Subjects	5
Personnel	5
Setting	5
Experimental Design and Evaluation	7
General Procedures	7
Selection and Definition of the Response	7
Selection and Definition of the Reinforcers	8
Token Exchange	8
Reinforcement Procedures	10
Priming	10
Results	11
Discussion	13
EXPERIMENT II	
Subjects	17
Personnel	17
Setting	17
Testing Under Reinforcement	17
Results	18
Discussion	20
EXPERIMENT III	
Subjects	21
Setting	21
Personnel	21
Experimental Design	21
Preparations for Testing Under Reinforcement	22
Response Definition	22
Reinforcer Definition	22
Reinforcer Priming	22
Response Priming	23
Test Administration	23
Results	24
Discussion	29
CONCLUSIONS	30
REFERENCES	32

LIST OF TABLES

Number		Page
1	Age, I.Q., and *Metropolitan Readiness Test* Score for the 24 Subjects	6
2	A Summary of Items Available for Token Exchange and Cost Per Item	9
3	Results From the *Metropolitan Readiness Test* Of the Six-Week Token Program	12
4	*Metropolitan Readiness Test* Scores for Each Subject When Tested Before and After Six Months of Standard Classroom Procedures	14
5	Per Cent of Change Between Pretest and Posttest Scores From the *Metropolitan Readiness Test*	15
6	*Metropolitan Readiness Test* Scores Under Standard and Reinforcement Conditions Before and After Introduction of Reinforcement Into Test Setting	19
7	Statistical Comparison of Subtests Scores on the *Metropolitan Achievement Test*	25
8	Comparison of Pretest and Posttest Raw Scores on the *Metropolitan Achievement Test* Obtained Under Standard Conditions	26
9	Compariosn of Raw Scores on the *Metropolitan Achievement Test* Obtained Under Standard Conditions with Scores Obtained Under Maximal Reinforcement Conditions	27
10	Grade Equivalents (Average Subtests) for May, 1970 on the *Metropolitan Achievement Test* Under Standard Conditions	28

ABSTRACT

In an attempt to evaluate the relative effectiveness of a token program in a school for the trainable mentally retarded, twenty-four children were matched into two groups. One group entered a six-week program where academic performance was reinforced with a wide range of incentives, while the second group continued in the standard teaching program of the school. At the beginning and end of the study a standardized test was given to both groups. Statistical treatment of the data showed significant improvement on the test for the group in the reinforcement program while the control group showed no improvement during the six weeks of standard classwork. Further, it was found that when a standardized test was administered under two motivational conditions, both groups showed improvement in score when tested with consequences of maximal reinforcement for correct performance. When this experimental probe of testing under reinforcement conditions was replicated with a normal population, the results showed statistically significant increases in scores achieved with maximal reinforcement consequences for correct performance.

INTRODUCTION

Educational programs for the trainable mentally retarded (TMR) child vary widely: some communities have provided special classrooms in the regular school for this severely mentally retarded group (I.Q. range is generally 30 to 50) while other areas have set aside special schools or institutions for their TMR populations (about two per thousand of school age population). Another alternative has been for these children to remain at home often with no educational attention. The decision regarding the placement of these children and the provision of tax monies for the necessary facilities has often varied as a function of the reports of investigators on the effectiveness of these various approaches.

Research in the area has generally reported little, if any, effective changes in the trainable population under a number of techniques and in a variety of settings. Goldstein (1956) studied data from a two-year project involving 173 children in 22 TMR classes. As tested by standard psychometric tools, he found no indications of mental growth over the two-year period and only 7% improved to the degree that they were admitted to educable mentally retarded (EMR) classes. Johnson and Capobianco (1957) observed 17 classes in a variety of settings in a two-year study. They found no significant difference among the groups (half-day, full-day school sessions, institutionalized) in social behavior or in language and stated that no major changes were produced along any of the dimensions measured. Similarly, Hottel (1958) found no significant changes over a one-year period between day classes training and parental training at home. The one study with positive results reports only partial success: Peck (1960) found that school training made a difference in learning progress as compared between these groups.

One difficulty in obtaining significant results may be the lack of effective evaulative instruments; intelligence tests have generally been devised to measure children over the age of five. The trainable mentally retarded child falls below this category and few tests are designed to measure this mental level. A deficit of trained teachers and controlled experiments might also be cited as contributors to the paucity of information regarding the relative

effects of various programs. Another factor responsible for the failure to find any differences in the children's behavior may well be the training methods themselves. Indeed, in these studies the time has come to focus attention toward other methods and techniques rather than varying components of the older, ineffective ones.

Learning theory has contributed a number of basic concepts to predict and effect changes in behavior. Learning itself has been defined as "a change in the probability of a response" (Skinner, 1968) but this simple statement presents a paradigm of the basic tenets of learning theory. Every organism has a repertoire of responses in interaction with his environment. There is however, a great variation as to the frequency and strength of particular responses since these variables are a function of the events which follow a particular response. If a certain behavior, such as a child smiling, is followed by another behavior which is pleasant and pleasing, such as a smile from another person, then the child will exhibit a smile with greater frequency. If, however, his smile is met with a blank face or is ignored the smile will tend to decrease in frequency.

Many more complex behaviors have been built using the simple, clear-cut model. The field of applied behavior analysis has extended this general approach to a wide variety of populations. For example, Lovaas, et al. (1965) reduced the self-destructive behavior of an autistic child systematically following the child's appropriate behaviors with social reinforcement. Bijou and Orlando (1961) used simple operant techniques to condition four retardates in a free-operant situation. Phillips (1968) applied these same principles to increase the appropriate social behavior of "pre-delinquent" boys and Ayllon and Azrin (1965) developed normal behaviors in adult schizophrenics through the use of these behavioral principles.

As an outgrowth of this basic application of learning principles a system has been developed to include a wide range of complex behaviors using "natural" events in the environment to generate and strengthen behavior (Ayllon and Azrin, 1968). This system consists of an economy where a special currency (tokens) is adopted that can be exchanged for a large number of "back-up"

incentives or reinforcers. Just as money in itself has no intrinsic value, so the tokens have no rewarding or reinforcing characteristics by themselves. Their value derives from exchange, thus allowing for different interests and appetites among individuals at various times. This eliminates the possibility of guessing what it is that motivates a given individual at a given time.

Classroom applications of the token system are particularly exciting. Disruptive behavior, a frequent complaint of teachers, has been successfully eliminated through the use of token reward for appropriate behavior (O'Leary and Becker, 1967 and O'Leary et al., 1969). Wolf et al. (1968) used grade points as tokens to raise the grade average of students in a remedial classroom. Birnbrauer et al. (1965) demonstrated the effectiveness of programmed instruction with a class of retarded children when that technology was combined with a motivational system based on token reinforcement. These studies represent a small sample of findings in the literature concerning extension of the token economy system. Data from these studies, coupled with the results from additional research conducted at Milton Avenue School by Ayllon, Schneider, and McCullen (1970) led to the design of a program which differed from the earlier programs using token reinforcement. Not only was the population at the Milton School severely retarded, but a standardized education test was used to evaluate the effectiveness of the token program.

Typically, the criterion for behavioral evaluation of reinforcement procedures is a direct measure of the frequency of the behavior in the presence and absence of reinforcement. In this study, however, while daily classwork was the behavior which was reinforced, the evaluation tool was a standardized test. Although this is not the usual procedure for behavioral studies, this design is followed by educators on every academic level. This procedure neither created a new measurement tool specifically tailored for material under reinforcement in this program, nor altered the teacher's program to conform to the objectives of the *Metropolitan Readiness Test* (MRT). Rather, the motivational system provided the teacher with a means to control the attention and concentration of the children for academic class work.

The question which initially began this study was: what effect would a motivational system for daily classwork have on a standardized school readiness test when no attempt is made to link these two academic activities together? To phrase this question another way, what is the effectiveness of a token reinforcement system on academic objectives when the form, as well as the sequence and frequency of presentation of academic subject matter, is left entirely to the teacher's discretion?

The objective of this program, then, was to compare the relative effects of a standard school instruction with a six-week token program on a group of TMR children. The tool for evaluation chosen was a pretest and posttest measure on the MRT.

EXPERIMENT I

Subjects

Twenty-four trainable mentally retarded children were the subjects of this study. There were 15 girls and 9 boys with an age range of 9 to 17 years (average was 12.6 years). The range of I.Q. scores was 33 to 56 points (average was 47.4). Prior to the six-week program the children were matched on the basis of their ages and scores on the *Metropolitan Readiness Test* (MRT) and assigned to either the experimental group or the control group. Table 1 shows the age, I.Q., and MRT score for each of the 24 children.

Personnel

The teacher and aide in this study were from the regular school staff. Each performed her regular duties and no extra training was needed to implement this program. Graduate students from Georgia State University carried out phases of the program which were conducted outside the classroom. All activities proceeded according to the standard schedule outlined for use in this school.

Setting

This study was conducted at Milton Avenue School, which is an urban school for the trainable mentally retarded. Facilities employed were one of the

TABLE 1

AGE, I.Q., AND *METROPOLITAN READINESS TEST* SCORE FOR THE 24 SUBJECTS

Subject	Age Years-Months	I.Q.	Score
Experimental			
1	9-10	48.0	53.0
3	10- 0	55.0	42.0
5	8- 9	52.0	39.0
7	12- 2	48.0	52.0
9	15- 7	41.0	60.0
11	17- 0	46.0	63.0
13	16- 1	46.0	53.0
15	16-11	46.0	45.0
17	16- 3	47.0	69.0
19	13- 6	51.0	62.0
21	14- 7	50.0	57.0
23	8- 2	48.0	35.0
Average	13- 2	48.2	52.5
Control			
2	9- 2	33.0	43.0
4	10- 8	56.0	46.0
6	10- 6	38.0	49.0
8	11- 9	50.0	44.0
10	14- 9	47.0	57.0
12	13- 3	48.0	56.0
14	12- 6	44.0	48.0
16	12- 3	51.0	52.0
18	14- 6	51.0	67.0
20	13- 8	51.0	62.0
22	10- 6	46.0	54.0
24	9- 7	46.0	34.0
Average	11- 9	46.8	51.0
Average for Both Groups	12- 6	47.4	51.75

regular classrooms, the school cafeteria, and the gymnasium. No special equipment was provided nor was any change made in the daily program for the rest of the classes in the school.

Experimental Design and Evaluation

The basic design utilized was the pretest and posttest comparison of two matched groups, one under experimental treatment and the other under the standard classroom procedures for this school. The MRT used in the evaluation is designed to measure language skill and aptitude, number knowledge, visual and motor skills, and the ability to follow directions. The data from these scores constitute the basis for evaluation and analysis of this program.

General Procedures

After all children had taken the initial MRT, each was assigned to a classroom which was conducted according to the usual procedure employed by the school. All 24 children shared the same educational program for six months which included the following activities: academic lessons, indoor and outdoor recreation, group music activities, homemaking training, and others. At the end of this period, the alternate form of the MRT was given to the children. On the basis of these scores and chronological age, 24 children from eight classrooms were matched into 12 pairs. Twelve of these children were assigned to one classroom under one teacher and were exposed to a six-week program of reinformcement for academic performance. The other 12 children remained in their original classrooms and continued under the same program with no changes in procedures. Both groups took the MRT together at the end of the six weeks.

Selection and Definition of the Response

Correct academic work including spelling, writing, copying, arithmetic, and reading was selected as the target behavior for reinforcement. Each day the teacher and the aide prepared a packet of academic work for each child according to his ability level. Each packet had an equal number of pages and the child received one token for each page of correct work.

Selection and Definition of the Reinforcers

Two opportunities were available each day for the exchange of the tokens earned for correct academic work. The first opportunity was in a special area of the gym which had been set aside for this program. An entertainment booth was set up at which the children could purchase a variety of items and activities. The children spent one-half hour each day in the entertainment area.

The second opportunity for token exchange was at lunchtime. All children in the school had a hot lunch each day which they ate in the school lunchroom. The experimental group, however, could pay admission with a token to a special dining area adjoining the lunchroom. This area had a larger table with padded chairs, carpeting, and was cooler and quieter than the lunchroom. Also, the children could purchase cokes and candy after finishing their lunch in the special dining area and could play records. A summary of items available for token exchange with the cost of each is given in Table 2.

Token Exchange

The tokens used were painted bottle caps which were both distinctive and inexpensive to prepare. They were lightweight, easy to handle, and durable, thus eliminating the need for replacement. As each child made his selection of exchange items for his tokens, he dropped the correct number of tokens into a special jar kept by the graduate student in charge of the implementation of exchange procedures. Colorful displays and charts showed the token prices for the various items and activities available in the entertainment area. The graduate student kept a record of each child's purchases per day.

Admission to the special dining area was also managed by the graduate student who collected a token from each child as he entered the area. Once inside, the child could watch television, play records, or buy candy and cokes by dropping a token into the special jar. The graduate student also kept a daily record of the purchases made by each child.

TABLE 2

A SUMMARY OF ITEMS AVAILABLE FOR TOKEN EXCHANGE AND COST PER ITEM

Areas		Cost in Tokens
Entertainment	Candy (2 M & M's plus Kool-Aid (1 paper cup)	1
	Balloons	2
	Riding in Kiddie Kars (5 min.)	1
	Renting skates (5 min.)	1
	Jewelry (bracelets, necklaces)	3-5
	Ribbons	2
	Perfume	3-5
	Toy cars	3-5
	Toy guns	3-5
	Airplanes (wooden glider type)	3
	Comic books	3 to buy, 1 to rent
	Play one record	1
	Play piano (5 min.)	1
	Rent Frisbee (5 min.)	1
	Play with large inner-tube (5 min.)	1
Lunch	Admission or re-admission to special dining area	1
	Play one record	1
	Watch television (5 min.)	1
	Coke (1 paper cup)	2
	Candy (2 M & M's)	1

Reinforcement Procedures

A token was delivered to the child contingent upon his correct completion of one page of academic material. The teacher and the aide delivered tokens for this written work which was immediately corrected as the child completed each page. Two opportunities were available when the child could earn tokens for academic work: during the period preceding the gym activities and also immediately before lunchtime.

The token exchange was set up so that no attempt was made to alter the teacher's daily schedule; therefore, the opportunity for earning tokens was limited to those times of the day explicitly set aside for academic work. The afternoon schedule did not allow for a regular period of academic work for earning tokens which would interfere with the teacher's procedures. However, the children had two opportunities per day to earn tokens, each of which was immediately followed by an opportunity to exchange the tokens earned.

Priming

In order to acquaint the children with token earning and token exchange, a priming procedure was used in the settings in which each of these behaviors were to occur. In the classroom, the teacher initially explained that the children would have an opportunity to earn tokens which could be exchanged for a variety of items which she enumerated. She then went on to describe the procedure for receiving tokens, that is, contingent upon the correct completion of each page of the work regularly given to them each day. As each token was delivered to the child, the teacher reiterated that it was for the correct work done by the child. When the children finished their work, they were instructed to take their tokens with them to the entertainment area in the gym. The graduate students encouraged the children to play with the games and toys. This procedure follows the "Reinforcer-Sampling Rule" (Ayllon and Azrin, 1968) which emphasized the importance of familiarizing the user with a certain event in the specific context in which it is to occur. This allows for the reinforcing properties of the event to be exhibited and experienced by the user. When it became apparent that certain items were

reinforcing to a child, as suggested by his continued attention to the item, the graduate student told the child to keep it (or to continue playing with it, as appropriate to the particular item or activity) for a certain number of tokens. The child was then instructed as to the procedure for token exchange.

A "free day" was made available in the special dining area in order to acquaint the children with this lunchtime option. They were treated to candy, cokes, and records. Thereafter, the procedures for admission and for acquiring the candy and cokes were followed as previously described.

Results

Children in the token reinforcement program scored significantly higher than the children under standard classroom instruction (t = 2.55, significant at .01 level). Table 3 gives the raw scores and per cent of change between the pretest and posttest of the MRT for the six-week token program.

Further, when a within-group comparison is made, that is, comparing the experimental group in terms of its pretest and posttest scores, and making a similar analysis for the control group, the results indicate that the experimental group averaged 3.75 points higher on the standard-procedure portion of the final evaluation than they had averaged on that same portion of the previous test (t = 3.04, significant beyond .05 level). On the other hand, the control group averaged 2.75 points lower than their previous test scores (t = 3.17 in a negative direction, significant beyond .01 level). The decrease in the scores for the control group is noteworthy as it indicates that this group, far from improving over the previous testing session, actually "lost ground" academically.

While this difference might be interpreted as due to the effect of reinforcement, another interpretation might be that what is most desirable about reinforcement is that its results can be effective when it is applied for short periods of time; possibly the same results could be attained, in a somewhat longer time frame, by standard classroom instruction. If this were the case, then the cost of reinforcement procedures may militate against the use of reinforcement

TABLE 3

RESULTS FROM THE *METROPOLITAN READINESS TEST*
OF THE SIX-WEEK TOKEN PROGRAM

Subjects	Pretest Score	Posttest Score	Change Between Scores	Per Cent Of Change
Experimental				
1	24.0	24.0	0	0.0
3	22.0	23.0	+ 1	4.5
5	18.0	20.0	+ 2	11.1
7	24.0	30.0	+ 6	25.0 #
9	27.0	31.0	+ 4	14.8
11	30.0	31.0	+ 1	3.3
13	26.0	30.0	+ 4	15.4
15	22.0	33.0	+ 11	50.0 #
17	32.0	39.0	+ 7	21.9
19	31.0	26.0	- 5	- 16.1
21	24.0	30.0	+ 6	25.0 #
23	15.0	22.0	+ 7	46.7 #
Average	24.58	28.25	+ 3.75	16.8
		t = 3.04 *		
Control				
2	19.0	10.0	- 9	- 47.3 #
4	19.0	18.0	- 1	- 5.3
6	21.0	14.0	- 7	- 33.3 #
8	18.0	22.0	+ 4	22.2
10	26.0	25.0	- 1	- 3.8
12	25.0	25.0	0	0.0
14	27.0	21.0	- 6	- 22.2
16	27.0	20.0	- 7	- 25.9 #
18	25.0	24.0	- 1	- 2.9
20	29.0	32.0	+ 3	10.3
22	24.0	23.0	- 1	- 4.2
24	25.0	18.0	- 7	- 28.0 #
Average	24.58	21.83	- 2.75	- 11.7
		t = - 3.17 **		

* Significant at .05 level.
** Significant at .01 level.

\# Subjects whose change between scores exceeded 25 per cent.

for comparable results would be attained with standard procedures more cheaply although in a somewhat longer time period.

To assess this possibility, an additional analysis was made of test data corresponding to six months of standard classroom instruction for the 12 children in the experimental group. Test data of the experimental group (Table 4) show that only two out of 12 children raised their scores by more than 25% while two children actually decreased their scores by the same margin. Further, when this same group was given six weeks of token reinforcement, their MRT pretest and posttest scores (Table 5) showed that four of the 12 children increased their scores beyond 25%.

A cursory look at Table 5 shows that six out of 12 children show a decrease in their score after six months of standard instruction. The token program, on the other hand, reduced this poor performance from six children to one child.

Increased time alone does not educate a child. Further, increased time not only is inerfective but it is also expensive. These results suggest that a six-week program such as the one reported here might have saved the State $280.55 (Georgia Department of Education,(1968), or the difference in cost between keeping a child in school for six months ($374) vs. for a six-week period ($93.45). If the objective of keeping a child in school is to teach him material that can be assessed through a standardized test, then the school did not accomplish this objective in six months while the six-week program did.

The results of this study clearly indicate that increased time (and, therefore, money) does not educate a child: rather, it is the application of well-based techniques and scientific principles that motivates a child to learn. True education is the result of the child's desire to learn.

Discussion

The results from Experiment I showed a significantly higher MRT score for the token reinforcement group as compared with the children who were

TABLE 4

METROPOLITAN READINESS TEST SCORES FOR EACH SUBJECT WHEN TESTED BEFORE AND AFTER SIX MONTHS OF STANDARD CLASSROOM PROCEDURES

Subject	First Test Score	Second Test Score	Per Cent Of Change
Experimental			
1	52.0	53.0	1.9
3	24.0	42.0	75.0 *
5	37.0	39.0	5.4
7	--	52.0	--
9	65.0	60.0	- 7.7
11	83.0	63.0	- 36.1
13	66.0	53.0	- 19.7
15	66.0	45.0	- 31.4
17	65.0	69.0	6.2
19	64.0	62.0	- 3.2
21	59.0	57.0	- 3.4
23	22.0	35.0	59.1 *
Control			
2	40.0	43.0	7.5
4	45.0	46.0	2.2
6	24.0	49.0	104.2 *
8	52.0	44.0	- 15.4
10	--	57.0	--
12	64.0	56.0	- 12.5
14	42.0	48.0	14.2
16	43.0	52.0	20.9
18	73.0	67.0	- 9.6
20	67.0	62.0	- 7.5
22	34.0	54.0	58.8 *
24	34.0	34.0	0.0
Total Averages	50.95	51.75	

* Subjects whose score exceeded 25 per cent increase.

TABLE 5

PER CENT OF CHANGE BETWEEN PRETEST AND POSTTEST SCORES FROM THE *METROPOLITAN READINESS TEST*

Subject	Before and After 6 Months of Standard Classroom Instruction	Before and After 6 Weeks of Token Reinforcement Procedures
Experimental		
1	+ 1.9	0.0
3	+ 75.0	+ 4.5
5	+ 5.4	+ 11.1
7	(pre score not taken)	+ 25.0 *
9	- 7.7	+ 14.8
11	- 36.1	+ 3.3
13	- 19.7	+ 15.4
15	- 31.4	+ 50.0 *
17	+ 6.2	+ 21.9
19	- 3.2	- 16.1
21	- 3.4	+ 25.0 *
23	+ 59.1	+ 46.7 *
Average	+ 4.2	+ 16.8

* Subjects whose change in score exceeded 25 per cent.

under standard classroom instruction. Since the comparison between the experimental and control groups is based on performance on a standardized test the question is what do the low scores for the control group actually reflect. Admittedly, test scores reflect a variety of conditions beyond simply "what the child knows": his health, mood, attitude, and the test items themselves combine to affect the final score to a great degree. One of the objectives of this project has been to find out what happens when one of these variables is manipulated, namely the child's motivation to perform well. Under standard conditions of testing, however, there is no means by which to determine exactly how accurately the test score reflects the child's knowledge, for these variables cannot be eliminated or even realistically controlled. Indeed, test scores often reflect poor academic skills, but they may also reflect lack of motivation to do well in the criterion test. Conceivably, some children in Experiment I learned much in the classroom under standard instruction, but were unwilling to make an effort during the criterion test. There is an inherent problem, then, in dealing with test scores and test score comparison, for this acknowledged variability presents a "natural" contamination of data in terms of detailed analysis. However, if motivation for performing during the criterion test is the key factor underlying the poor performance of the control group, this thesis can be checked empirically.

EXPERIMENT II

Experiment II was designed to determine the role of reinforcement on performance on the criterion test, the *Metropolitan Readiness Test* (MRT). Experiment I was conducted in the morning and Experiment II was conducted after the lunch hour.

All 24 children had been given the first portion of the *Metropolitan Readiness Test* (MRT) during the morning (Experiment I). In the afternoon session all 24 children were given the second portion of the MRT, thus ruling out the possibility of other experiences confounding the measures. In this manner, two measures could be analyzed as premeasure and postmeasure of the effects of reinforcement on test performance. The morning scores were

obtained using token reinforcement for correct test performance.

Subjects

The 24 children who participated in Experiment I, also participated in Experiment II.

Personnel

The teacher who had administered the first portion of the test was assisted by two graduate students in the administration of the second portion of the test.

Setting

In order to control for as many variables as possible which could affect the test performance, the setting chosen was the classroom in which the MRT testing had taken place earlier that day. The children had gone to lunch after the initial test session and were instructed to return to the same room.

Testing Under Reinforcement

The second part of the test was set up so that the children could earn tokens for correct performance on the test. They could then exchange the tokens for a wide variety of back-up items.

While the children were at lunch a large table was set up in a prominent place and was used to display toys, balloons, comics, puzzles, games, and edibles, such as candy, cokes, and ice cream bars. When the children were asked to return to the classroom many of them were reluctant, particularly those 12 who came from seven other classrooms. Apparently the prospect of further testing was not a welcome one. Upon seeing the display, however, and being told that everyone would have a chance to earn tokens to trade for these items, they became interested and asked what they had to do. The teacher then explained that they would be given another booklet just like the first one and that for every correct answer, they would earn a token

which they could exchange for any of the items on display. As an additional incentive, the children were told that the ten making the highest scores would also earn a dollar as well as being allowed to spend their tokens in the classroom.

In order to give the control group, who was unfamiliar with the token system, the experience of earning and exchanging tokens, tokens were given for correct performance on the first three sample items included in the booklets. The child was given a token for each correct answer. In addition, if he answered all three sample items correctly he was given a dime. At the end of these three sample items the children were given an opportunity to exchange their tokens for candy. When this priming procedure was completed the teacher began the first subtest. At the end of each subtest the graduate students assisted the teacher in correcting the booklets and giving the child one token per correct item. The number of correct items per subtest was marked on the front of the booklet to facilitate final totaling. When the testing was completed the total test scores for the 24 children were ranked and the top ten children were awarded their $1 prize. They were permitted to exchange their tokens first and the others followed.

Results

When children, whose school background did not include token reinforcement for performance, take a standardized test, they score higher under reinforcement than under standard testing conditions. This score increase also applies to the group familiar with reinforcement procedures. Out of a possible total of 51 points, the control group scored 21.83 and 28.08 on the two portions of the test for an average increase under reinforcement of 6.25 points. This is significant beyond the .01 level (t = 5.96). Further, for the control group the average increase in score under reinforcement was 34.6% with a range of 0.0% to 140.0%. The experimental group averaged 28.25 on the first part under no reinforcement and 35.42 on the second part of the test. This increase of 7.17 points is significant beyond the .01 level (t = 5.89). For the experimental group, the individual per cent of increase varied from 3% to 58% averaging an increase of 26.5%. Table 6 summarizes this data for all 24 subjects.

TABLE 6

METROPOLITAN READINESS TEST SCORES UNDER STANDARD AND REINFORCEMENT CONDITIONS BEFORE AND AFTER INTRODUCTION OF REINFORCEMENT INTO TEST SETTING

Subject	Standard Conditions	Reinforcement Conditions	Change Between Scores	Per Cent Of Change
Experimental				
1	24.0	38.0	+ 14	58.0
3	23.0	29.0	+ 6	26.0
5	20.0	24.0	+ 4	20.0
7	30.0	41.0	+ 11	37.0
9	31.0	39.0	+ 8	27.0
11	31.0	40.0	+ 9	29.0
13	30.0	38.0	+ 8	27.0
15	33.0	36.0	+ 3	9.0
17	39.0	41.0	+ 2	3.0
19	26.0	40.0	+ 14	54.0
21	30.0	33.0	+ 3	10.0
23	22.0	26.0	+ 4	18.2
Average	28.25	35.42	+ 7.17	26.5
		t = 5.89 **		
Control				
2	10.0	24.0	+ 14	140.0
4	18.0	23.0	+ 5	27.8
6	14.0	21.0	+ 7	50.0
8	22.0	28.0	+ 6	27.0
10	25.0	34.0	+ 9	36.0
12	25.0	33.0	+ 8	32.0
14	21.0	21.0	0	0.0
16	20.0	27.0	+ 7	35.0
18	34.0	40.0	+ 6	18.0
20	32.0	40.0	+ 8	25.0
22	23.0	26.0	+ 3	13.0
24	18.0	20.0	+ 2	11.0
Average	21.83	28.08	+ 6.25	34.6
		t = 5.96 **		

** Significant at .01 level.

The only difference between the two testing conditions was the motivation to perform well in the second test situation. In effect, there was a differential consequence for doing a test item correctly, namely the token which could be exchanged later for candy and toys.

The effect of token reinforcement was uniform across both groups regardless of the academic and reinforcement history of the child. Eight of the 12 children in the control (67%) increased their scores by more than 25%. Similarly, seven of the 12 children in the experimental group, when tested under reinforcement conditions, increased their scores by more than 25%.

Discussion

When one considers the limited repertoire of the trainable mentally retarded child, the increase in scores under one session of exposure to reinforcement is indeed impressive. The child is motivated to do well on the test for there is a differential consequence for his efforts. In effect, merely going through the motions of taking the test is insufficient to obtain token reinforcement.

These results give support to the view that the test score of special populations, like the disadvantaged, the retarded, and the handicapped, are often the results of motivation during test and not a reflection of the extent of their skills.

EXPERIMENT III

Significant gains were achieved when reinforcement principles were applied in an academic setting, even when the program was of a short duration (Experiment I). Furthermore, significant increases in test scores were also produced in a single exposure to reinforcement (Experiment II). Although these data are impressive and statistically beyond chance occurrence, the choice of subject population, namely the trainable retarded, limits its usefulness. Indeed, the improvement in performance of the trainable retarded may be possible because their skills and behavior repertoire, in general, are not as complex

as those of normal children. Normal children, on the other hand, are probably operating at the level of behavioral complexity least affected by incentives such as those used with the retarded. Fortunately, this is a question that can be empirically tested. The purpose of Experiment III, then, was to assess the effect of one exposure to maximal reinforcement conditions on a standardized test when a normal population was used. In effect, this was a replication of Experiment II with normal children.

Subjects

Thirty fourth grade pupils were the subjects of this exploratory study. All had been enrolled in Jessie Mae Jones School since the beginning of the school year. All subjects were black. Their I.Q. scores ranged from 73 to 138 (mean was 92.8).

Setting

The experimental phase of this study took place in the classroom regularly occupied by this class.

Personnel

Three graduate students from Georgia State University administered the procedures under the direct supervision of Dr. T. Ayllon. The teacher was also present but did not participate in the proceedings.

Experimental Design

Two test sessions provided the pretest and posttest data for analysis in this study. The measurement tool for this probe was the *Metropolitan Achievement Test* (MAT) Elementary Battery. Statistical treatment was applied to the group data from the first test which was taken under standard conditions and the second test administration under maximal reinforcement conditions. Scores on each of the seven subtests were tallied for all subjects and comparisons of these scores were made with the parallel portion of the other test data.

Preparations for Testing under Reinforcement

Since this particular achievement test is recommended by the testing manual to be administered in 3 to 5 different sittings, it was necessary to select only portions of the entire battery. It was decided to administer the second half of each subtest for the test session under reinforcement. Booklets were prepared for each child with that portion of the test marked and clearly delineated.

Response Definition

The behavior selected as the response for which the children would receive reinforcement was correct performance of the items on the (MAT) Elementary Battery. For each item that was done correctly, they would receive one token. The areas covered by the seven subtests were word knowledge, word discrimination, reading, spelling, language, arithmetic computation, and arithmetic concepts.

Reinforcer Definition

A large colorful display was prepared showing toys, puzzles, books, ribbons, candy, games, and other items. In addition, cokes and ice cream were also available. To each of these items was attached a card which gave the value or cost of the item. This number represented the number of tokens needed by the child to purchase the item. The ten highest scorers on the test would earn one dollar to spend as they chose.

Reinforcer Priming

In order to familiarize children with the items available for exchange of tokens earned on the test, a priming procedure was employed. The children were encouraged to examine and handle the reinforcers on display and to learn the "prices," or the number of tokens they would need to purchase the items. Further, the children were given candy and told that they could earn more candy as well as their choice among the back-up items in exchange for the tokens they would earn for correct performance.

Response Priming

While interest in earning tokens was at its peak, a response priming session was employed. This was designed to acquaint the children with the method of earning the tokens by their correct performance. The first three items of the first subtest were selected as sample items. The children were told that for each of these items which they did correctly, they would earn a token; if all three were done correctly they would not only have their 3 tokens to exchange for candy but would earn a dime in addition.

After each item was completed, the work was checked by the graduate students and, if correct, the child was praised and given his token. After all three items had been completed and corrected, the children were invited to exchange the tokens for a variety of candies. In addition, those earning all three tokens for the sample items were given a bonus of one dime.

Test Administration

After priming procedures were completed, the test procedures were explained to the class. One graduate student pointed out that one-half of the first subtest was marked in yellow and the other half was not. The items in yellow (the second half) were those which when done correctly, would earn tokens for the child. The other half would earn no tokens, even if done correctly. It was further explained that there was only enough time to do one part and the children had to choose which they would like to do. Each subtest was preceded by a reminder to the children that they were free to choose which of the items they wished to work on.

After the first subtest, the graduate students scored each child's work and gave him a token for each correct item. This same procedure was followed for each subtest. At the end of the test session, the scores were added and the ten highest scorers were announced and awarded their one-dollar prize. Each child was then given an opportunity to exchange his tokens for his choices of the back-up reinforcers.

Results

When normal children are given a choice between answering test items that result in incentives and those that do not, they choose to work on the items that result in reinforcement. Furthermore, under maximal reinforcement conditions the children increased their scores significantly beyond the scores obtained on the pretest.

It was found that on 5 of the 7 subtests, there was a significant increase between the standard vs. the reinforcement conditions (Tables 7 through 10). Further, on the overall test, the increase as measured by a t test showed that the effect of reinforcement was significant beyond the .01 level (t = 5.90). These findings support the work of Ayllon, Kelly, and McCullen (1970), which demonstrated the effect of reinforcement on test performance. Although the children reached a high level of achievement on the first test, they were not at their peak performance level as indicated by the tremendous score increase when the variable of motivation was introduced.

One might argue that although the reinforcement is a powerful experimental method, it still cannot compare with the effects of long-term student-teacher interaction in the classroom. The focus of the teacher's efforts are on those skills considered vital for every child; the standardized test which is designed to measure these skills, then should reflect the teacher's success in meeting her own goals in these academic areas. Thus, it was decided to explore the measured effects of classroom experience with the results of test administration under reinforcement conditions.

The MAT had been administered at the beginning of the school year and again after a period of 8 months. During this interval, standard classroom procedures were in use by the teacher. When the change between these two test administrations, both of which were taken under standard conditions, are analyzed by a t test, 6 of the 7 subtests showed a significant increase at the .01 level with an overall t value of 7.10. Thus, this class had made much improvement over the 8-month period as reflected in the standardized test measures. Table 7 summarizes the statistical data for each of these score comparisons as well as the analysis of the test scores taken under reinforcement conditions.

TABLE 7

STATISTICAL COMPARISON OF SUBTEST SCORES
ON THE *METROPOLITAN ACHIEVEMENT TEST*

	Tests Under Standard Conditions Before and After 8 Months		Tests Under Standard Conditions and Under Reinforcement Conditions	
Subtest	*t* Score	Level of Significance	*t* Score	Level of Significance
Word Knowledge	2.03	n.s.	4.16	.01
Word Discrimination	2.29	.05	0.75	n.s.
Reading	3.09	.01	- 0.82	n.s.
Spelling	7.98	.01	3.52	.01
Language (A + B)	3.89	.01	3.13	.01
Arithmetic Computation	4.07	.01	2.58	.02
Arithmetic Concepts	6.32	.01	2.36	.05
TOTAL	7.10	.01	5.90	.01

TABLE 8

COMPARISON OF PRETEST AND POSTTEST RAW SCORES ON THE *METROPOLITAN ACHIEVEMENT TEST* OBTAINED UNDER STANDARD CONDITIONS

	Word Know.		Word Discrim.		Reading		Spelling		Lang. (A+B)		Arith.Comp.		Arith.Conct.		TOTAL	
Subject	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1	23	19	15	18	11	19	14	17	21	24	18	23	16	17	118	137
2	3	8	4	10	3	15	9	17	11	20	7	10	4	8	41	88
3	5	5	3	6	6	7	0	16	11	15	1	5	2	1	28	55
4	9	7	9	14	4	15	11	16	12	19	2	0	1	4	48	72
5	6	4	7	3	9	1	1	10	12	13	10	14	1	6	46	51
6	7	9	3	12	2	8	1	13	11	18	1	7	1	3	26	70
7	3	7	11	13	2	6	3	12	11	15	8	10	1	3	39	66
8	0	1	2	7	3	4	0	12	0	3	5	16	0	2	10	45
9	6	8	4	8	3	8	3	14	10	12	1	7	0	2	27	59
10	4	9	3	9	4	13	5	15	14	15	2	9	4	2	36	72
11	2	4	4	4	1	5	0	4	8	14	7	11	1	5	23	47
12	4	7	6	7	1	3	0	5	10	11	1	5	2	3	24	41
13	7	6	1	5	6	3	0	5	13	7	6	10	1	2	34	38
14	6	6	2	4	4	9	0	8	4	16	2	9	2	1	20	53
15	5	4	6	8	4	3	0	10	9	13	4	6	0	2	28	46
16	1	7	6	6	1	8	0	3	8	12	2	4	1	0	19	40
17	7	13	2	9	5	7	1	9	4	3	3	4	1	4	23	49
18	4	5	5	4	5	4	0	2	9	9	0	2	0	1	23	27
19	9	6	4	5	3	4	1	7	6	10	1	4	1	2	25	38
20	2	4	1	6	3	3	0	6	8	17	0	3	1	3	15	47
21	6	4	3	7	2	1	0	1	12	5	0	2	1	1	24	21
22	4	8	6	2	0	6	0	9	10	11	1	5	1	2	22	43
23	5	3	0	1	5	6	0	2	6	11	8	6	2	3	26	32
24	8	4	4	6	6	2	0	5	6	14	0	3	1	2	25	36
25	6	8	5	8	2	6	0	0	13	9	9	1	2	3	37	35
26	6	1	5	1	5	2	0	3	7	14	1	4	1	4	25	29
27	4	5	4	8	4	4	0	8	8	7	3	3	1	2	24	37
28	0	2	0	3	0	5	0	1	1	8	1	2	0	0	2	21
29	6	6	4	10	8	8	0	0	12	9	0	2	2	0	32	35
30	4	4	2	5	4	7	0	0	7	15	0	0	2	0	19	31
	$\Sigma d = 22$		$\Sigma d = 78$		$\Sigma d = 76$		$\Sigma d = 181$		$\Sigma d = 91$		$\Sigma d = 88$		$\Sigma d = 32$		$\Sigma d = 572$	
	$\bar{d} = 0.73$		$\bar{d} = 2.60$		$\bar{d} = 2.53$		$\bar{d} = 6.03$		$\bar{d} = 3.10$		$\bar{d} = 2.93$		$\bar{d} = 1.07$		$\bar{d} = 17.06$	
	$\Sigma d^2 = 274$		$\Sigma d^2 = 490$		$\Sigma d^2 = 776$		$\Sigma d^2 = 1,589$		$\Sigma d^2 = 829$		$\Sigma d^2 = 610$		$\Sigma d^2 = 124$		$\Sigma d^2 = 15,740$	

Note: Pretest administered in September, 1969.
Posttest administered in May, 1970 (after 8 months of standard academic class work).

TABLE 9

COMPARISON OF RAW SCORES ON THE *METROPOLITAN ACHIEVEMENT TEST* OBTAINED UNDER STANDARD CONDITIONS WITH SCORES OBTAINED UNDER MAXIMAL REINFORCEMENT CONDITIONS

	Word Know.		Word Discrim.		Reading		Spelling		Lang. (A+B)		Arith.Comp.		Arith.Conct.		TOTAL	
Subject	$Post_1$	$Post_2$	$Post_1$	$Post_2$	$Post_1$	$Post_2$	$Post_1$	$Post_2$	$Post_1$	$Post_2$	$Post_1$	$Post_2$	$Post_1$	$Post_2$	$Post_1$	$Post_2$
1	19	22	18	17	19	19	17	20	24	25	23	18	17	17	137	138
2	8	11	10	11	15	10	17	16	20	23	10	16	8	5	88	92
3	5	8	6	11	7	7	16	19	15	12	5	8	1	7	55	72
4	7	20	14	14	15	6	16	15	19	17	0	6	1	5	72	83
5	4	3	3	8	1	10	10	13	13	20	14	16	6	6	51	76
6	9	14	12	13	8	6	13	14	18	17	7	9	3	4	70	77
7	7	10	13	10	6	3	12	16	15	18	10	14	3	6	66	77
8	1	6	7	7	4	5	12	9	3	18	16	16	2	4	45	65
9	8	12	8	9	8	5	14	15	12	14	7	8	2	1	59	64
10	9	9	9	8	13	7	15	13	15	18	9	9	2	12	72	76
11	4	8	4	5	5	9	4	6	14	17	11	11	5	6	47	62
12	7	9	7	8	3	14	5	14	11	12	5	9	3	0	41	66
13	6	9	5	3	3	6	5	7	7	17	10	10	2	2	38	54
14	6	9	4	3	9	8	8	12	16	15	9	6	1	5	53	58
15	4	6	8	10	3	3	10	12	13	21	6	5	2	2	46	59
16	7	6	6	8	8	5	3	12	12	15	4	3	0	5	40	54
17	13	8	9	9	7	3	9	15	3	12	4	4	4	3	49	54
18	5	10	4	8	4	6	2	6	9	12	2	2	1	1	27	45
19	6	3	5	5	4	5	7	7	10	14	4	10	2	1	38	45
20	4	9	6	3	3	4	6	4	17	17	8	10	3	1	47	48
21	4	7	7	6	1	4	1	2	5	15	2	7	1	2	21	43
22	8	8	2	6	6	7	9	11	11	5	5	3	2	3	43	43
23	3	6	1	7	6	2	2	3	11	12	6	7	3	2	32	39
24	4	4	6	3	2	4	5	6	14	15	3	11	2	0	36	43
25	8	7	8	4	6	8	0	0	9	14	1	3	3	3	35	39
26	1	7	1	3	2	7	3	9	14	10	4	4	4	1	29	41
27	5	8	8	5	4	5	8	6	7	8	3	3	2	3	37	38
28	2	7	3	5	5	6	1	6	8	10	2	2	0	0	21	36
29	6	7	10	3	8	4	0	0	9	11	2	1	0	0	35	26
30	4	5	5	7	7	2	0	0	15	10	0	0	0	0	31	24
	$\Sigma d = 74$		$\Sigma d = 10$		$\Sigma d = -2$		$\Sigma d = 58$		$\Sigma d = 75$		$\Sigma d = 39$		$\Sigma d = 22$		$\Sigma d = 276$	
	$\bar{d} = 2.47$		$\bar{d} = .33$		$\bar{d} = -0.07$		$\bar{d} = 1.93$		$\bar{d} = 2.50$		$\bar{d} = 1.30$		$\bar{d} = 0.73$		$\bar{d} = 9.20$	
	$\Sigma d^2 = 490$		$\Sigma d^2 = 251$		$\Sigma d^2 = 526$		$\Sigma d^2 = 378$		$\Sigma d^2 = 824$		$\Sigma d^2 = 272$		$\Sigma d^2 = 250$		$\Sigma d^2 = 4,656$	

Note: $Post_1$ administered under standard conditions (May, 1970).

$Post_2$ administered under maximal reinforcement conditions (two weeks after $Post_1$).

TABLE 10

GRADE EQUIVALENTS (AVERAGE SUBTESTS) FOR MAY, 1970
ON THE *METROPOLITAN ACHIEVEMENT TEST*
UNDER STANDARD CONDITIONS

$Post_1$ Grade Equivalent	Subject	$Post_1$ Grade Equivalent	Subject
6.74	1	3.53	14
4.81	2	3.45	19
4.53	4	3.44	20
4.37	6	3.39	22
4.23	10	3.37	13
4.21	7	3.26	27
3.78	5	3.13	24
3.74	15	3.13	23
3.73	11	3.09	26
3.71	16	2.94	18
3.70	3	2.73	29
3.67	8	2.55	21
3.60	12	2.47	28
3.60	9	2.43	25
3.59	17	2.11	30

Average = 3.58

Note: $Post_2$ grade equivalents are not calculable since the test under reinforcement was on the second portion of each subtest only. Conversion tables included with the administration manual do not allow for these kinds of score alterations with regard to standard score and grade equivalent extractions from the raw scores.

In order to determine the overall stability of the test scores over the time period of all three test administrations, the data were given a one-way analysis of variance (ANOVA) with repeated measures. This test yielded an F ratio of 79.80, significant well beyond the .001 level. This means that the amount of change from each test score to the following test score was significant.

A further probe was then initiated to determine if each of the three test means were significantly different statistically from the other two test scores. The Neuman-Keuls method using a modified q statistic was used in this analysis. The q_r value of the means of the test scores taken before and after the 8-month period of standard classroom conditions was $q_r = 11.82$ (significant beyond the .01 level). The difference between the second test score and the test score taken two weeks later under reinforcement conditions was $q_r = 5.702$, also significant beyond the .01 level. This statistical treatment demonstrated that the effect of testing under maximal reinforcement conditions was powerful enough to match the academic gains made in 8 months or 34 weeks of classroom instruction and interaction.

Another means of determining the relative effect of reinforcement procedures is to compute the per cent of change between each pair of test scores. Almost two-thirds of the children raised their test scores by more than 50% following 8 months of classroom activities. Moreover, 21 of the 30 children raised their test scores by more than 25% after 34 weeks in the classroom. Only two weeks later, however, this high scoring level was surpassed by more than 40% of this group: 8 children increased their scores between 25% and 50% over the previous test score while 4 children exceeded a 50% increase when taking the test under maximal reinforcement conditions. The effect of the motivating properties of the reinforcement procedures then was to produce performance beyond an already significantly high score with the addition of a single powerful variable.

Discussion

The purpose of this experimental probe was to determine the relative effects of reinforcing correct performance on test scores. The results of

this test administration were compared with scores from the same test administered under standard testing conditions.

Statistical analysis showed that significant improvement did take place over the 8 month interval between the first two test sessions, both across the subtests and on the overall score. When reinforcement in the form of tokens was given for correct performance, scores increased significantly even beyond the already high level of performance achieved by this class. Prior to the test administration the class of 30 students used in this experiment averaged 3.59 overall grade equivalent which is above the overall grade equivalent for the fourth grade at Jessie Mae Jones School. Although these children appeared to be functioning at their best, their enhanced performance under maximal motivation reflects greater academic potential.

The conclusions from this data cannot be lightly dismissed as hypothetical or highly speculative. The rate of change when the test was given under maximal reinforcement conditions matched the amount of change which accrued between test sessions when 8 full months, or 34 weeks, of standard educational procedures were in operation. In terms of cost alone, these data cannot be regarded lightly. When a one-session experiment can match the results of 8 months of academic teaching, these experimental procedures demand further investigation.

CONCLUSIONS

Experiment I showed a significant difference in the effect of two kinds of treatment on a standardized test score of children in the program. Experiment II showed that a significant increase in test scores is possible by the use of reinforcement even when the exposure is restricted to one session. Experiment III replicated the results of the second experiment by extending the data to validation on a normal population. These results, obtained from both a population typically limited in skills and ability as well as from a normal group of children, raise the question as to how much more could be gained from an entire school year's experience in this type of program. The

possibilities are endless and there are no limits to the extent of application of a program built along these dimensions. To create for a child an environment which rewards his efforts to learn in ways meaningful to him, not by prediction or guesswork but by testing out these consequences, is to make a place where the child's limits are not man-made or environmentally predetermined. It is the professional and moral duty of the educator to design the educational environment in such a way that the intellectual activity of the child is nurtured rather than stiffled.

REFERENCES

Ayllon, T. "Intensive treatment of psychotic behavior by stimulus satiation and food reinforcement," *Behavior Research and Therapy*, 1963, 1:53-61.

Ayllon, T. and Azrin, N. H. "The measurement and reinforcement of behavior of psychotics," *Journal of Experimental Analysis of Behavior*, 1965, 8:357-383.

Ayllon, T. and Azrin, N. H. *The Token Economy: A Motivational System for Therapy and Rehabilitation.* New York: Appleton-Century-Crofts, 1968.

Bijou, S. W. and Orlando, R. "Rapid development of multiple-schedule performances with retarded children," *Journal of the Experimental Analysis of Behavior*, 1961, 4:7-16.

Birnbrauer, J. S., Bijou, S. W., Wolf, M. M., and Kidder, J. D. "Programmed instruction in the classroom." In L. P. Ullman and R. Krasner (Eds.) *Case Studies in Behavior Modification.* New York: Holt, Rhinhart, and Winston, Inc., 1965. Pp.358-363.

Georgia Department of Education. *Cost per ADA Child*, 1968, Atlanta, Georgia.

Goldstein, H. *Report Number Two on Study Projects for Trainable Mentally Retarded Handicapped Children.* Issued by Vernon L. Nickell, Superintendent of Public Instruction. Springfield, Illinois, 1956.

Hottel, J. *An Evaluation of Tennessee's Day Class Program for Severely Mentally Retarded Trainable Children.* Nashville, Tennessee: State Department of Education, 1958.

Johnson, G. O. and Capobianco, R. J. *Research Project on Severely Retarded Children.* Albany, New York: Interdepartmental Health Resources Board, 1957.

Lovaas, O. I., Freitag, G., Gold, Vivian J., and Kassorla, Irene C. "Experimental studies in childhood schizophrenia: Analysis of self-destructive behavior," *Journal of Experimental Child Psychology*, 1965, 2:67-84.

O'Leary, K. D. and Becker, W. C. "Behavior modification of an adjustment class: A token reinforcement program," *Exceptional Children*, 1967, 33:637-642.

O'Leary, K. D., Becker, W. C., Evans, M.B., and Saudergas, R. A. "A token reinforcement in a public school: A replication and systematic analysis," *Journal of Applied Behavior Analysis*, 1969, 2:3-14.

Peck, J. R. A Comparative Investigation of the Social and Learning Adjustment of Trainable Children in Public School Facilities, Segregated Community Centers, and State Residential Centers. U. S. Office of Education Cooperative Research Program, Project No. SAE 6430, 1960.

Phillips, E. L. "Achievement place: Token reinforcement procedures in a home-style rehabilitation setting for pre-delinquent boys," *Journal of Applied Behavior Analysis*, 1968, 1:213-224.

Skinner, B. F. "The experimental analysis of behavior." In A. Charles Catania (Ed.) *Contemporary Research in Operant Behavior*. Atlanta: Scott, Foresman, and Company, 1968.

Wolf, M. M., Giles, D. K., Hall, V. R. "Experiments with token reinforcement in a remedial classroom," *Behavior Research and Therapy*, 1968, 6:51-64.